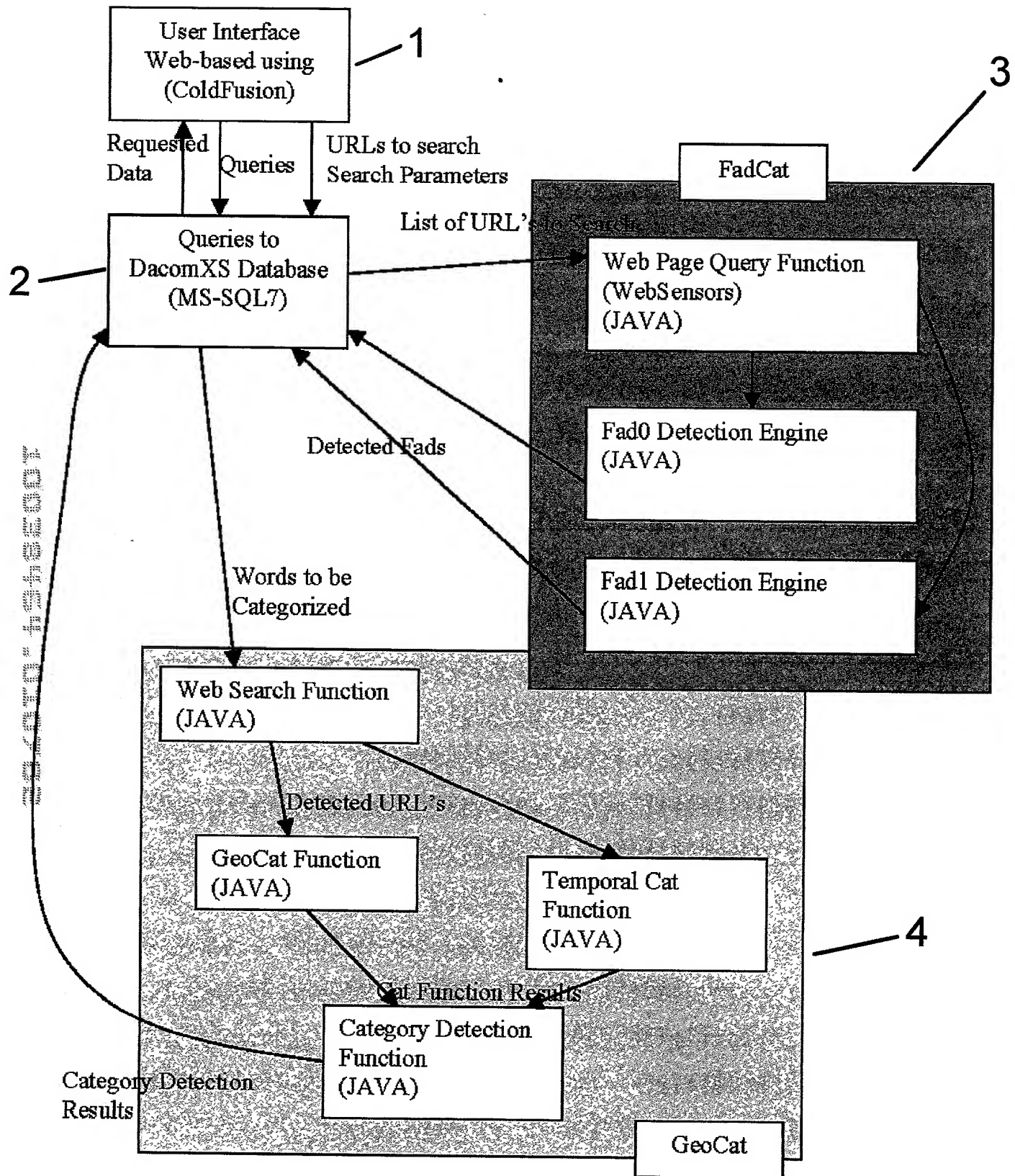
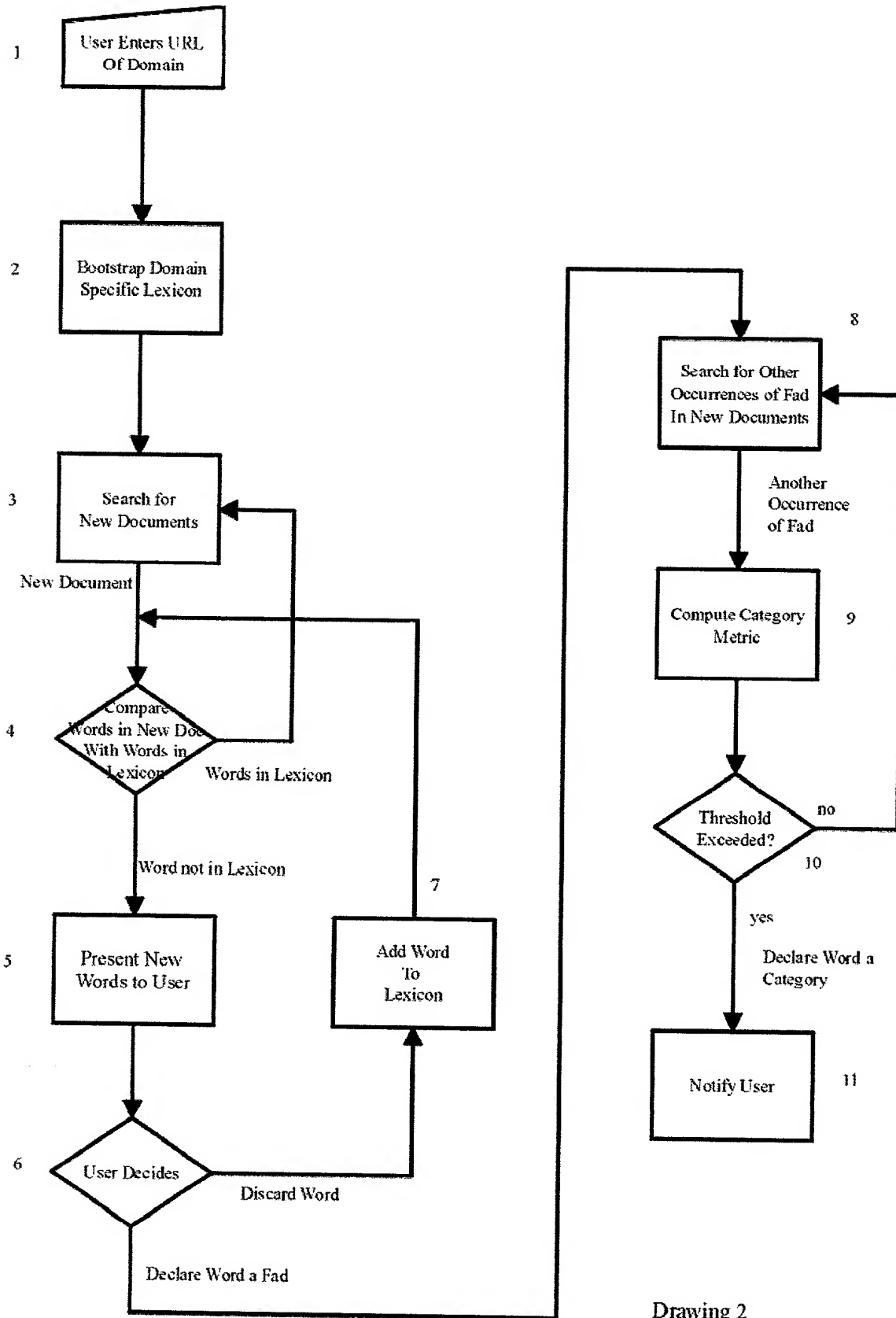


DRAWINGS

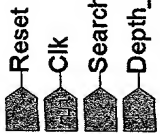


DRAWING 1



Drawing 2

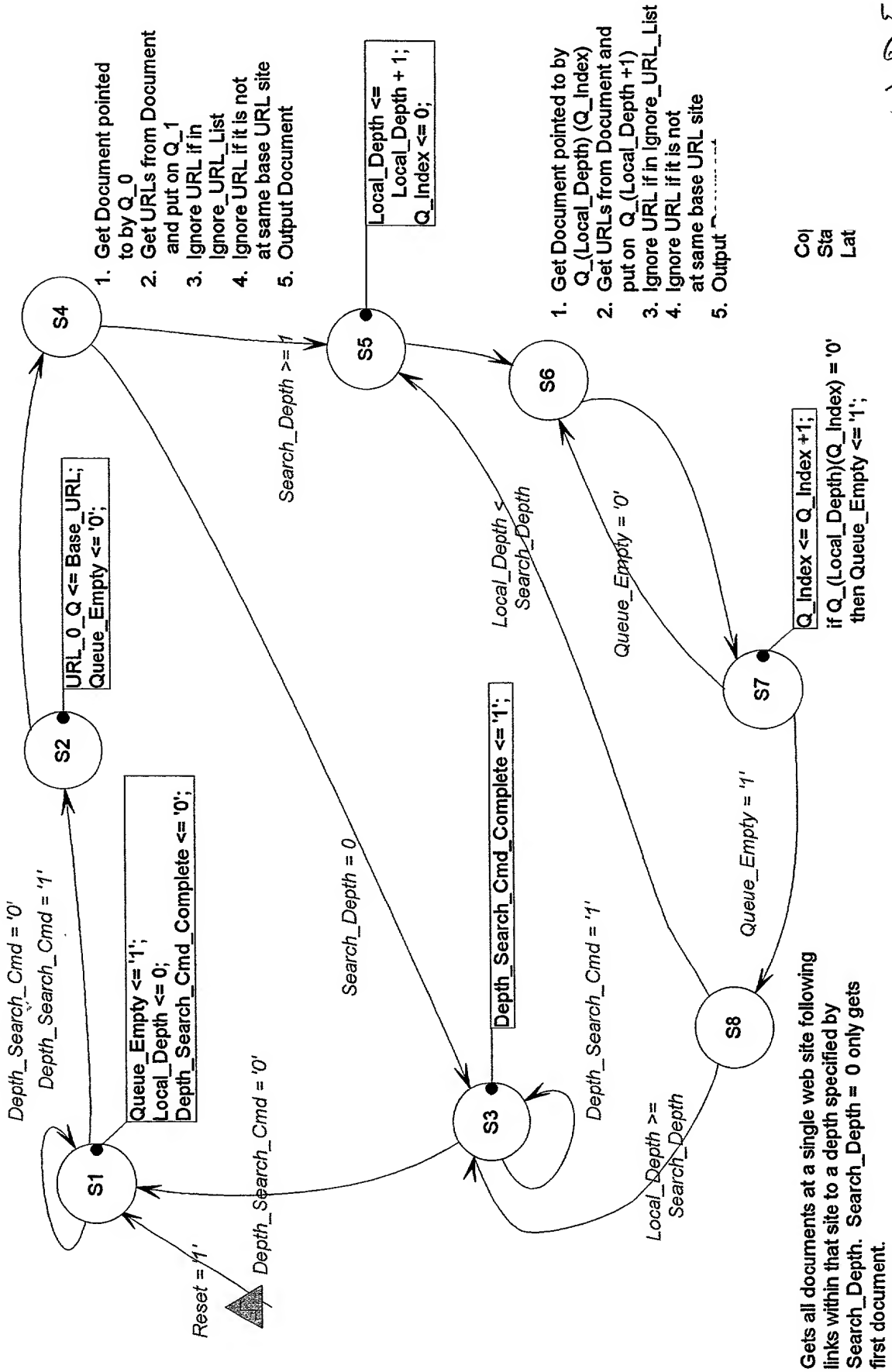
Get Docs To Depth



Ignore_URL_List[7:0]
Base_URL
Queue_Empty
Local_Depth

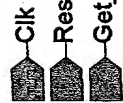
URL_0_Q
URL_1_Q[3:0]
URL_2_Q[3:0]
URL_n_Q[3:0]

Depth_Search_Cmd_Complete
Web_Site_Docs[63:0]
Q_Index



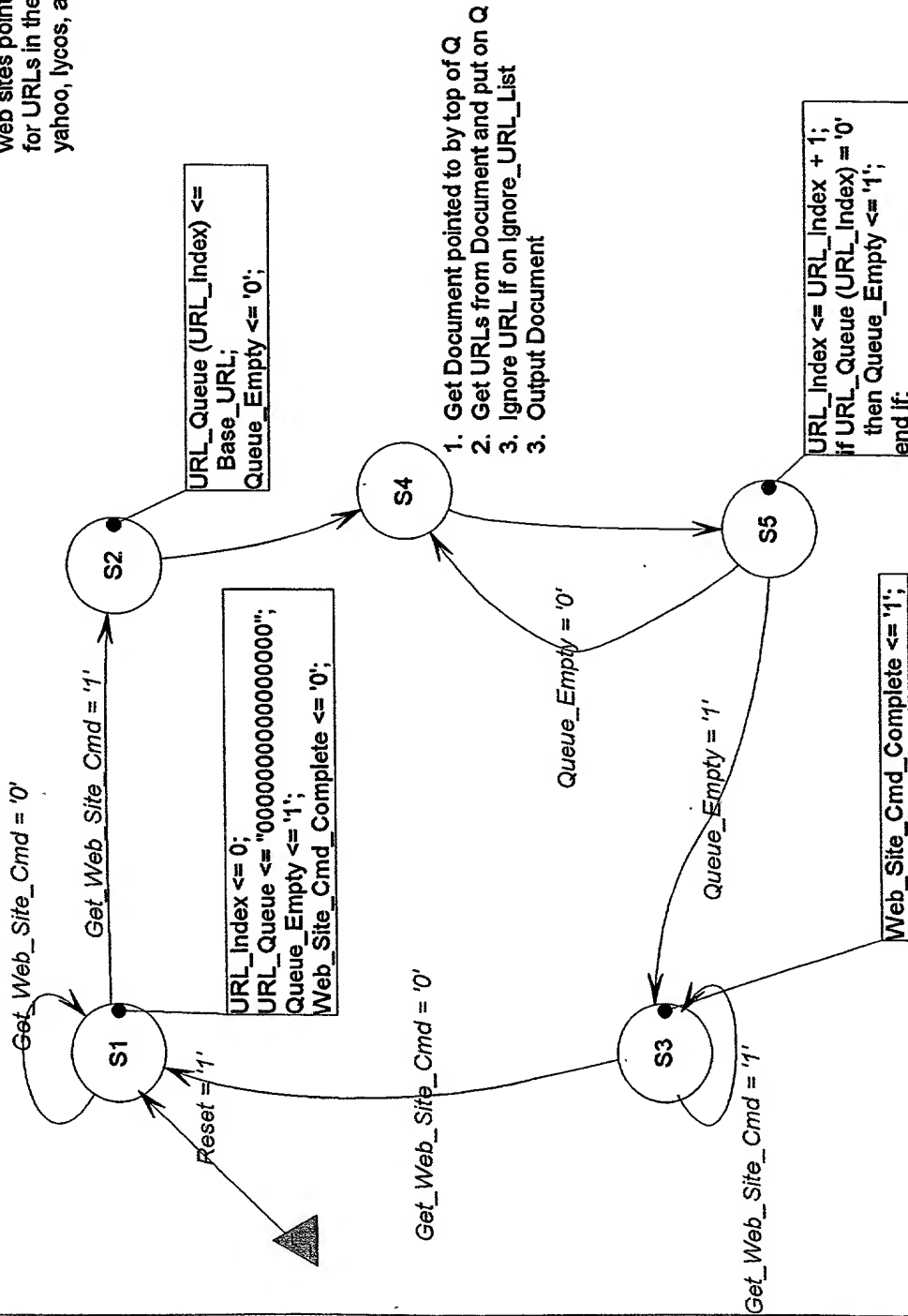
DRAWING 3

Get_All_Docs



Sreg0

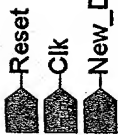
Gets all documents from a web site and from all web sites pointed to by that web site except for URLs in the Ignore_URL_List (sites like yahoo, lycos, altavista, etc.)



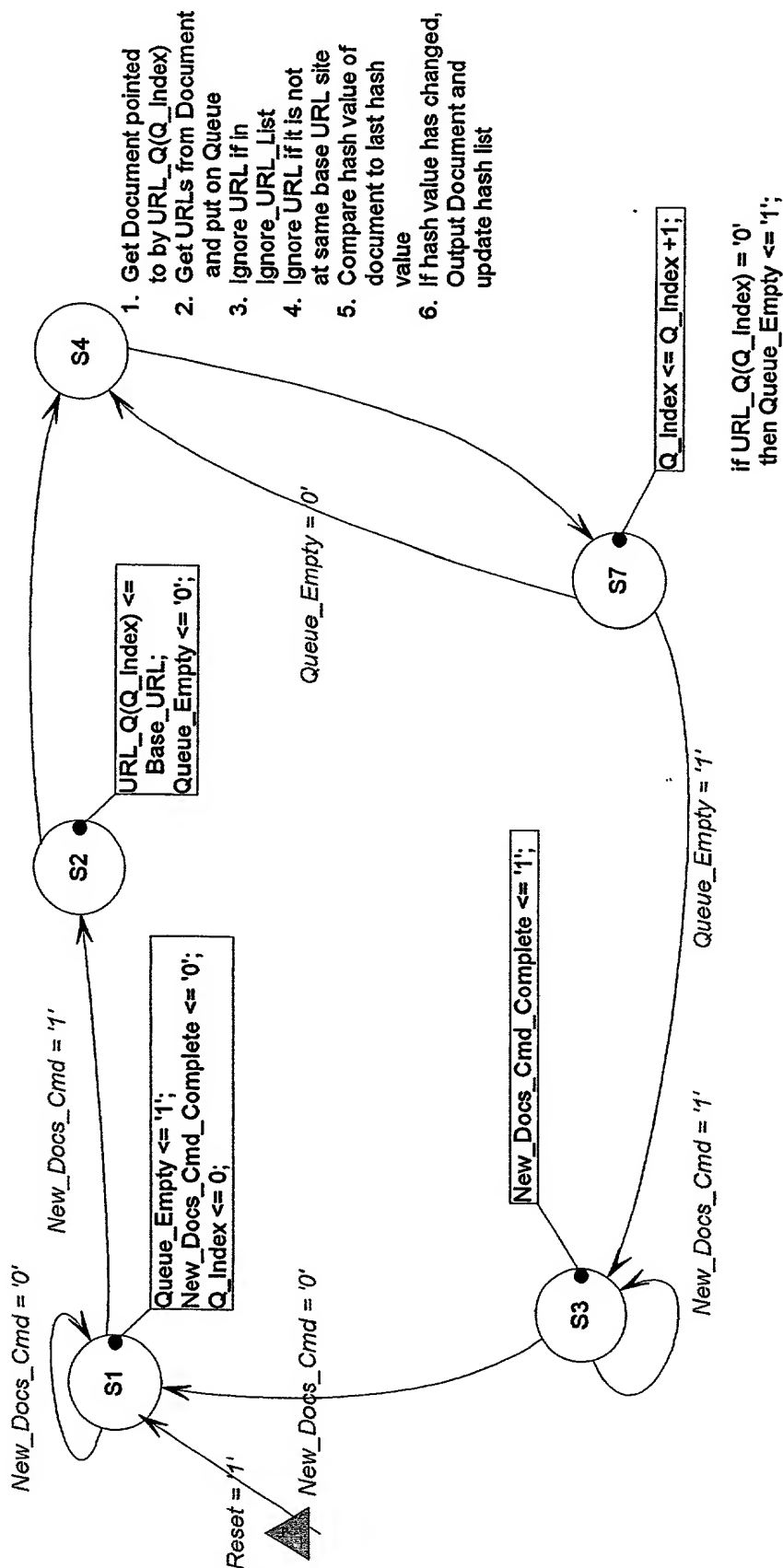
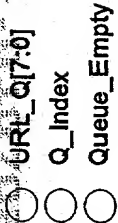
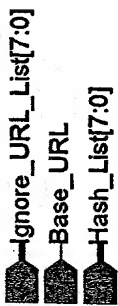
DRAWING 4

Copyright: K. J. Hintz, 2000
Start Date: 4/3/00
Latest Update: 4/4/00

46ED



Get_New_Docs



Gets only those documents from one site which have changed since the last time they were accessed.
Uses a list of document "hash" values stored by URL for later comparison

Copyright: K. J. Hintz, 2000
Start Date: 4/3/00
Latest Update: 4/4/00

DRAWING 5

5020

Fad_Det_0_FSM

Norm_0_Voc[15:0]

Fad_0_Voc[7:0]

Doc_Rdy

Fad_0_Det

Fad_0_New_Data

Fad_0_New

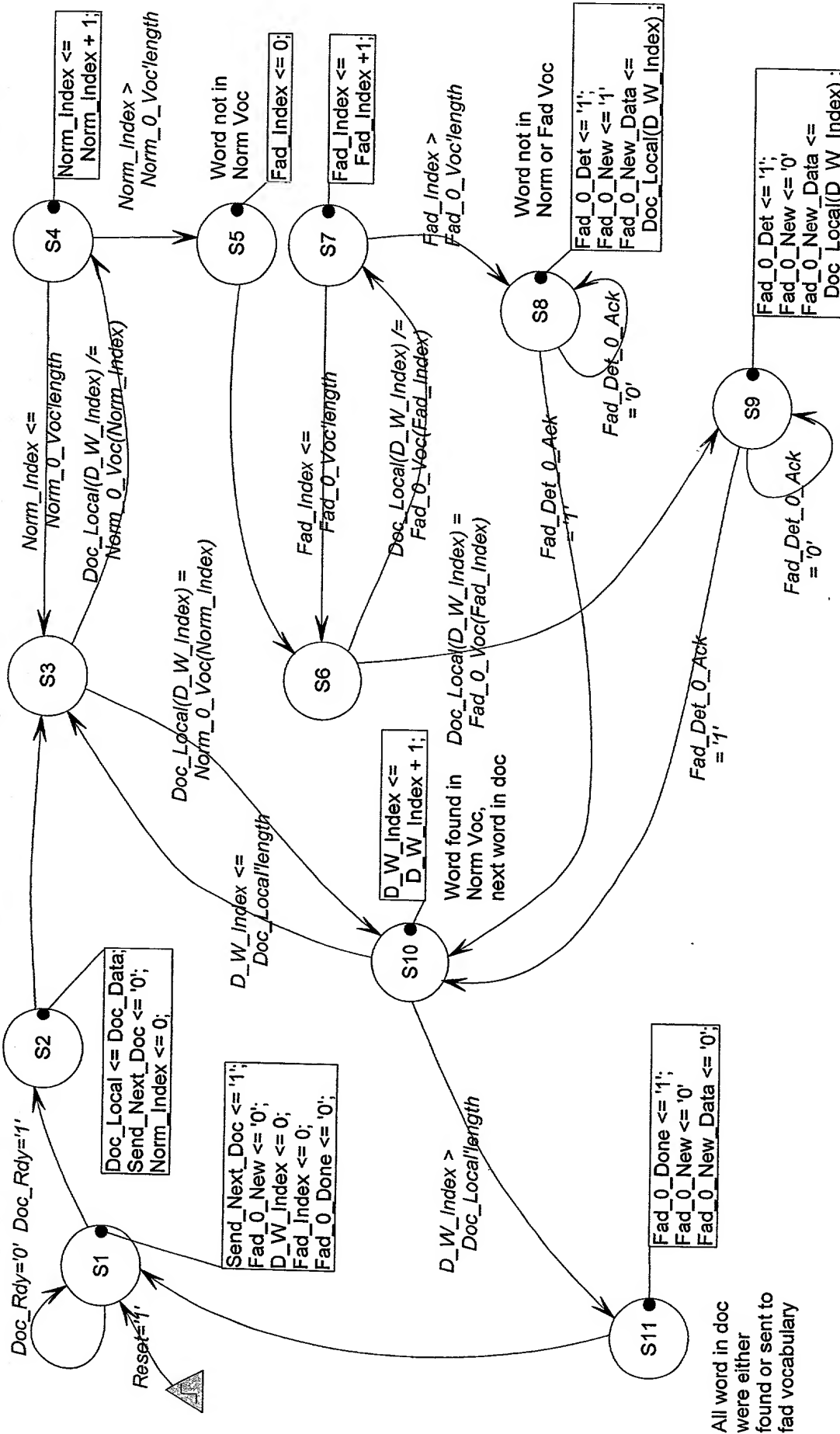
Fad_Index

Doc_Local[31:0]

Norm_Index

Fad_Det_0_Ac

Doc_Data[31:0]



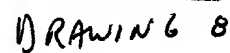
Copyright: K. J. Hintz, 2000
Start Date: 3/21/00
Latest Update: 3/22/00

DRAWING 6

6.6.3.0



Norm_1_Lex: strings of 2 words with "whitespace" separator replaced by one space only



Copyright: K. J. Hintz, 2000
Start Date: 8/1/00
Latest Update: 8/9/00

168